Cell fueling and metabolic energy conservation in synthetic cells

Sikkema, Hendrik R; Gaastra, Bauke F; Pols, Tjeerd; Poolman, Bert

Published in:
ChemBioChem

DOI:
10.1002/cbic.201900398

Publication date:
2019

A EUROPEAN JOURNAL OF CHEMICAL BIOLOGY

CHEMBIOCHEM

SYNTHETIC BIOLOGY & BIO-NANOTECHNOLOGY

Accepted Article

Title: Cell fueling and metabolic energy conservation in synthetic cells

Authors: Hendrik R Sikkema, Bauke F Gaastra, Tjeerd Pols, and Bert Poolman

To be cited as: *ChemBioChem* 10.1002/cbic.201900398

Link to VoR: http://dx.doi.org/10.1002/cbic.201900398

A Journal of

ChemPubSoc Europe

WILEY-VCH

www.chembiochem.org

Cell fueling and metabolic energy conservation in synthetic cells

Hendrik R. Sikkema[a], Bauke F. Gaastra[a], Tjeerd Pols[a] and Bert Poolman*[a]

Abstract: We aim for a blue print for synthesizing (moderately complex) subcellular systems from molecular components and ultimately for constructing life. Without comprehensive instructions and design principles we rely on simple reaction routes to operate the essential functions of life. The first forms of synthetic life will not make every building block for polymers de novo via complex pathways, rather they will be fed with amino acids, fatty acids and nucleotides. Controlled energy supply is crucial for any synthetic cell, no matter how complex. Here, we describe the simplest pathways for efficient generation of ATP and electrochemical ion gradients. We estimated the demand for ATP by polymer synthesis and maintenance processes in small cell-like systems, and we describe circuits to control the needs for ATP. We also present fluorescence-based sensors for pH, ionic strength, excluded volume, ATP/ADP, and viscosity, which allow monitoring and tuning of the major physicochemical conditions inside cells.

1. Introduction

> *"Life is not just about replication; it is also a coupling of chemical reactions – exergonic ones that release energy and endergonic ones that utilise it, preventing the dissipation of energy as heat".* [1]

"What is life" is one of the most intriguing and difficult questions to answer, even at the cellular scale. At the molecular level, however, it is well established that life is a system of self-sustained chemical processes. Biochemical networks direct cell growth and division, and through the uptake of nutrients, the conservation of metabolic energy and the excretion of waste, they maintain a dynamic state far from thermodynamic equilibrium. Other features of life-like systems are that they are kinetically controlled (orchestrated through feedback loops), self-organized and compartmentalized, which enables active, adaptive and autonomous behavior. Such properties are present even in the simplest forms of life.

The prospect of creating synthetic life has inspired people for many years. The Venter Institute, for instance, has demonstrated that a *de novo* synthesized genome containing less than 500 genes can lead to viable cells.[2,3] While creating a reduced cell by selectively removing components from a wild-type genome is an impressive achievement, this top-down approach leads to a minimal cell with a reduced set of biomolecules, but it does not reveal how the remaining gene products act together to create life, nor does it capture the links between metabolism, compartmentalization and the information contained in DNA. As a result, it has not yet been possible to rationally design and construct, using a bottom-up constructive approach, a simple form of life based on a limited number of molecular building blocks (see *e.g.* ref [4]). While our fundamental understanding of the individual building blocks of life is rapidly growing, putting a minimal set of components together such that life-like properties emerge remains a formidable, yet exciting challenge.

In our view, true understanding of "molecular life" requires the design and synthesis from scratch of systems with increasing complexity. This bottom-up assembly using molecular components has been referred to as synthetic biochemistry.[5] Fostered by the fields of biophysics and biochemistry and the need for quantitative studies of molecular building blocks, there has been rapid progress in the reconstitution and quantitative understanding of complex biological systems and processes, such as: complex membranes and transport systems[6], sophisticated DNA processing machineries[7,8], complex cytoskeletal systems[9], self-organized spatial protein patterns[10] and cell-free gene expression[11]. In addition, the possibilities for genome engineering have exploded with the development of powerful DNA assembly methods and the CRISPR.[12,13]

In the first part of this paper, we focus on the construction of cell-like systems from molecular building blocks, that is, the assembly and engineering of the components that enable a cell-like system to form ATP and generate electrochemical ion gradients and achieve energy homeostasis. This is one of the crucial networks that is essential for any life-like system as cells need both chemical and electrochemical fuel to enable endergonic reactions to occur. We describe systems already pioneered but also propose alternative pathways for metabolic energy conservation on the basis of known strategies employed by simple microbes. In the second part, we quantify the amount of ATP needed for a (minimal) synthetic cell to reproduce itself, while maintaining the same concentration of biomolecules in mother and daughter cells. We find that the majority of metabolic energy of our model cell is needed for protein synthesis and maintenance processes. In the third part we describe vesicle-based systems to encapsulate the metabolic networks for energy conservation, and the real-time monitoring of the internal conditions by fluorescence-based sensors. We also indicate where hurdles are expected in the construction of ever more complex systems.

1.1. Coupling of exergonic and endergonic reactions and measure of energy status

All known forms of life use two forms of energy currency: ATP and electrochemical ion gradients. The amount of free energy released upon hydrolysis of ATP to ADP plus inorganic phosphate is the same as that of other nucleoside triphosphates such as GTP, CTP, UTP or TTP, but ATP (and to a lesser extent GTP) is predominantly used when chemical energy needs to be coupled to endergonic reactions or processes (*i.e.* to shift the equilibrium). The energy stored in ATP is given by the phosphorylation potential (ΔG_p or $\Delta G_p/F$):

$$\Delta G_p = \Delta G^{0\prime} + 2.3RT log \frac{[ADP][Pi]}{[ATP]} \quad \text{(kJ/mol)} \quad \text{Eq. 1}$$

$$or \ \frac{\Delta G_p}{F} = \frac{\Delta G^{0\prime}}{F} + \frac{2.3RT}{F} log \frac{[ADP][Pi]}{[ATP]} \quad \text{(mV)}$$

Similarly, electrochemical proton or sodium ion gradients are most often used to drive membrane-bound processes, even though other types of ion and solute gradients exist. The F_0F_1-ATP synthase/hydrolase interconverts the free energy of the phosphorylation potential into an electrochemical proton gradient, hereafter referred to as proton motive force (Δp):

[a] Hendrik R. Sikkema, Bauke F. Gaastra, Tjeerd Pols & Bert Poolman
Department of Biochemistry
University of Groningen
Nijenborgh 4, 9747 AG Groningen, The Netherlands
E-mail: b.poolman@rug.nl

$$\Delta p = \Delta\Psi + \frac{2.3R}{F} log \frac{[H^+]_{in}}{[H^+]_{out}} = \Delta\Psi - Z\Delta pH \quad (mV) \quad Eq.\ 2$$

where 2.3*RT*/*F* equals 58 mV (at T=298 K) and is abbreviated as Z; *F* is the Faraday constant, *R* the gas constant and *T* is the absolute temperature. $\Delta G^{0'}$ = -30.5 kJ/mol, and typically ΔG_p ranges from -50 to -65 kJ/mol (or $\Delta G_p/F$ varies from -520 to -670 mV). A sodium motive force (Δs) can be formed in a similar manner:

$$\Delta s = \Delta\Psi + \frac{2.3RT}{F} log \frac{[Na^+]_{in}}{[Na^+]_{out}} = \Delta\Psi - Z\Delta pNa \quad (mV) \quad Eq.\ 3$$

From a control perspective it can be desirable to connect ATP and ion fluxes through a single enzyme such as F_0F_1-ATP synthase/hydrolase, but there are no fundamental principles that prohibit the two forms of energy to be formed and regulated independent of each other. As far as we are aware there are no known free-living forms of life without ATP synthase/hydrolase, but a few bacterial obligate endosymbionts lack the enzyme complex[14] and rely on substrate-level phosphorylation for their ATP production.[15]

2. Cell fueling systems

Respiratory organisms use the F_0F_1-ATP synthase to form ATP, whereas fermentative bacteria use the enzyme to hydrolyse part of their ATP obtained in catabolic reactions to generate an electrochemical ion gradient. At thermodynamic equilibrium, the phosphorylation potential equals the proton motive force times the number of protons (*n*) translocated per ATP. In formula:

$$\frac{\Delta G_p}{F} = n\Delta p \quad Eq.\ 4$$

This number is determined by the *c*-ring stoichiometry of ATP synthase/hydrolase and varies from 2.7 to 5, depending on the specific enzyme.[16] Some organisms exploit an F_0F_1-ATP synthase/hydrolase that translocates sodium ions instead of protons, hence the formation or utilization of a sodium motive force (Δs). In addition, most forms of life exploit so-called sodium-proton antiporters to interconvert Δp and Δs.

The F_1F_O-ATP synthase complex is one of the engineering masterpieces in the cell. We briefly discuss two important aspects of the complex, first the c-ring stoichiometry and second the regulation. The architecture of the c-ring, that is, specifically the copy-number of the c-subunit differs per organism from 8 copies for bovine mitochondria[17] to 15 copies in *Spirulina platensis*.[18] This leads to different proton-to-ATP ratios (Eq. 4). From an engineering point of view the high-speed gear (low copy-number) works well in organisms that are continuously exposed to a high proton motive force, like in the bovine mitochondria. A high copy number leads to a high torque gear, essential when the proton motive force is low, or variable.[19]

Because the magnitude of the Δp and ΔG_p varies and a cell needs both forms of metabolic energy above some threshold value, it is important to have regulation in place to restrict the directionality of operation. An important regulator of the bacterial F_1F_O-ATP synthase complex is the ε subunit. Structural data for this domain exists for two distinct conformations in different organisms.[20,21] Tsunoda et al.[22] have used cross-links to trap the ε subunit in both of these conformations in *E. coli*. They have then shown that in one conformation the synthase works in both directions, whereas in the other conformation the synthesis of ATP remains functional but the ATP hydrolysis is inhibited. Meyrat and von Ballmoos[23] have shown that high ATP/ADP ratios inhibit the ATP synthesis, preventing the proton motive force to be drained completely. These two regulatory mechanisms prevent futile cycling of the ATPase in either direction.

In heterotrophs, the oxidation of organic carbon yields CO_2 plus reducing equivalents such as NADH and $FADH_2$. The subsequent oxidation of NADH and $FADH_2$ results in the formation of an electrochemical proton gradient by the respiratory chain. The usage of the Δp by the F_0F_1-ATP synthase results in the synthesis of ATP, and the overall process is known as oxidative phosphorylation. This route to Δp and ATP formation is complex and requires numerous enzymes and cofactors. Nature offers alternative mechanisms to conserve metabolic energy through simple metabolic conversion (deamination of amino acids, oxidation of carboxylic acids) or the use of light. In the following sections we discuss a number of alternatives to oxidative phosphorylation for the synthesis of ATP. We focus on simple systems to ease application in synthetic cells.

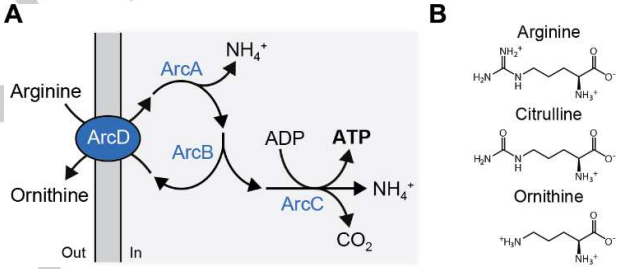

Figure 1. Arginine breakdown pathway Metabolic energy conservation by breakdown of arginine. (**A**) Schematic of the arginine breakdown pathway. ArcA, arginine deiminase; ArcB, ornithine transcarbamylase; ArcC, carbmate kinase; ArcD, arginine/ornithine antiporter. For every molecule of arginine imported, one molecule of ATP is produced, while the product ornithine is exchanged for arginine; NH_3 (formed from NH_4^+) and CO_2 diffuse out passively. (**B**) Structures of arginine, citrulline and ornithine at pH 7.

2.1. Arginine breakdown pathway

Deamination of arginine yields citrulline plus NH_4^+, which is catalyzed by the enzyme arginine deiminase. Subsequent phosphorolysis of citrulline by ornithine carbamoyltransferase yields ornithine plus carbamoyl phosphate, a reaction that is thermodynamically unfavorable (K_{eq} ~ 10^{-5}) but proceeds when the reaction products are drained. Carbamate kinase converts carbamoyl phosphate plus ADP into CO_2, NH_4^+ and ATP (Fig. 1A) and thereby conserves a large fraction of energy dissipated in the breakdown of the amino acid. Since the substrate arginine and product ornithine are structurally related (Fig. 1B), they can be transported by one and the same protein via a so-called antiport mechanism.[24] This property of coupling substrate and product fluxes is also possible in many other pathways and aids in keeping the reaction networks away from equilibrium. The arginine

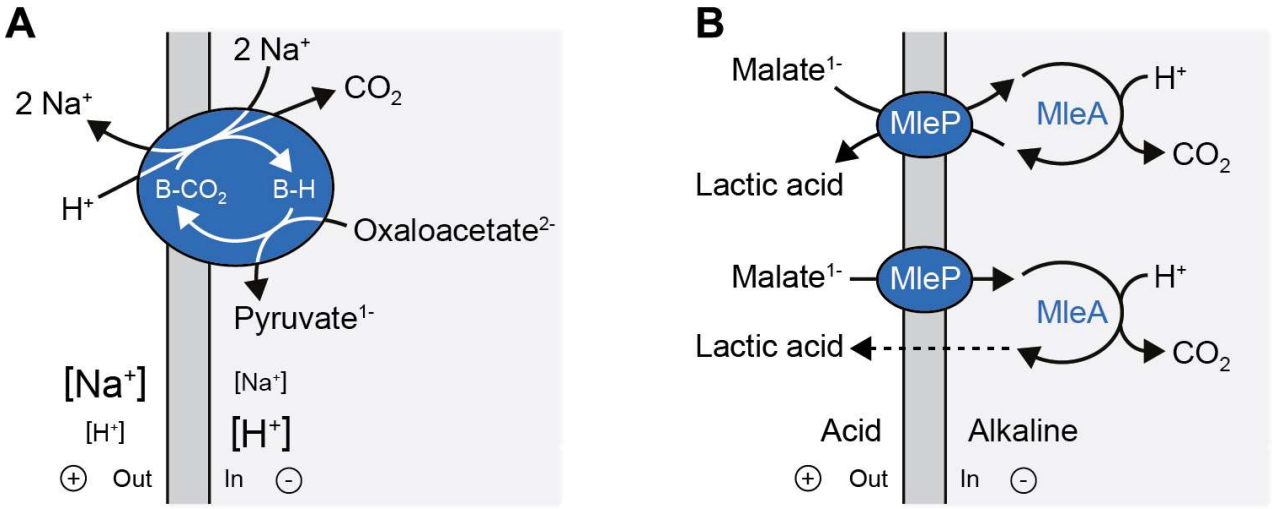

Figure 2. Decarboxylation pathways Metabolic energy conservation by decarboxylation of carboxylic acids (and amino acids, see Table I). (**A**) Schematic of the oxaloacetate decarboxylase Na^+ pump. For every molecule of oxaloacetate converted into pyruvate, 2 Na^+ ions are pumped out, while one H^+ is imported. The system thus generates an electrochemical sodium gradient (ΔΨ plus ΔpNa) and in theory a pH gradient inside acid relative to the outside. Since the outside volume is typically very large, the inverse ΔpH will only be formed if the cell density is high and the external buffering capacity is low. B, Biotin. (**B**) Schematic of the malolactic fermentation pathway. The decarboxylation of malate^{1-} consumes a H^+, while the product, lactic acid, can be either exchanged for malate^{1-} (top) or diffuse passively across the membrane (bottom). Both malate^{1-}/lactic acid exchange (top) and malate^{1-} uniport (with lactic acid diffusion) (bottom) generate a ΔΨ (inside negative relative to the outside) and ΔpH (inside alkaline relative to the outside).

breakdown pathway has been reconstituted in liposomes with ATP/ADP and pH sensors (Box I) in the vesicle lumen to report the synthesis of ATP and to monitor the changes in internal pH.[25] The system can sustain a constant level of ATP for many hours even when the load on the system is varied by the consumption of ATP for the uptake of solutes.

The overall reaction equation indicates that protons are consumed in the breakdown of arginine but in the vesicle system the actual internal pH is determined by (i) the rate of ATP production and consumption; (ii) the relative flux through the entire pathway and a futile route leading to citrulline; (iii) the diffusion of NH_3 out of the cell, leaving a proton behind for every NH_4^+ produced; and (iv) the fate of CO_2.

Ad (i) The synthesis of ATP is given by $ADP^{3-} + HPO_4^{2-} + H^+ \rightarrow ATP^{4-} + H_2O$. Thus, a proton is consumed in the synthesis and produced in the hydrolysis of ATP.

Ad (ii) The antiporter is not entirely specific for ornithine but also exchanges arginine for citrulline (not shown in the figure), creating a futile deamination route through the action of ArcA and ArcD.

Ad (iii) NH_3 can leave the vesicles by passive diffusion, which will leave a proton behind; the base/conjugated acid reaction of ammonia ($NH_4^+ \leftrightarrow NH_3 + H^+$; pK_A of 9.1) is fast.

Ad (iv) CO_2 can leave the vesicles by passive diffusion, but a high concentration of inorganic phosphate allows the formation of HCO_3^- and a proton, even in the absence of carbonic anhydrase.

Because the import of arginine and efflux of ornithine are coupled and NH_3 and CO_2 can diffuse out, the membrane-reconstituted arginine breakdown pathway constitutes an open system that enables long-term synthesis of ATP. A similar pathway can be envisaged in vesicles by employing the enzymes that convert agmatine into putrescine, CO_2 plus $2NH_4^+$, which also yields one ATP per substrate metabolized.

2.2. Decarboxylation pathways

The free energy released in the decarboxylation of dicarboxylic acids and amino acids is around -20 kJ/mol (Table I),[40] which is too little to directly make ATP from ADP plus inorganic phosphate (*vide supra*). The free energy change of a decarboxylation reaction can be stored in the form of an electrochemical ion gradient, which subsequently can be used to synthesize ATP

Table 1. Overview of decarboxylation systems. Antiport refers to the exchange of the indicated substrate and product. The net predominant charge of the molecules at pH 7 is indicated.

Substrate	***Product***	***Transport mechanism***	***Reference***
Malonate^{2-}	Acetate^{1-}	Electrogenic Na^+ pump	Berg et al 1997[26]
Oxaloacetate^{2-}	Pyruvate^{1-}	Electrogenic Na^+ pump	Dimroth, 1982[27]
Succinate^{2-}	Propionate^{1-}	Electrogenic Na^+ pump	Hilpert et al 1984[28]
Oxalate^{2-}	Formate^{1-}	Antiport	Anantharam et al 1989 [29]
			Hirai et al 2002[30]
Malate^{2-}	Lactate^{1-}	Antiport	Poolman et al, 1991[31]
		H-Malate$^-$ Uniport + lactic acid diffusion	Salema et al 1994; 1996[32,33]
Arginine^{1+}	Agmatine^{2+}	Antiport	Ilgü et al 2016[34]
Glutamate^{1-}	γ-amino butyric acid^0	Antiport	Richard et al 2004[35]
			Ma et al 2012[36]
Histidine^0	Histamine^{1+}	Antiport	Molenaar et al, 1993[37]
Lysine^{1+}	Cadaverine^{2+}	Antiport	Romano et al 2013[38]
Ornithine^{1+}	Putrescine^{2+}	Antiport	Romano et al 2013[38]
Tyrosine^0	Tyramine^{1+}	Antiport	Coton et al 2011[39]

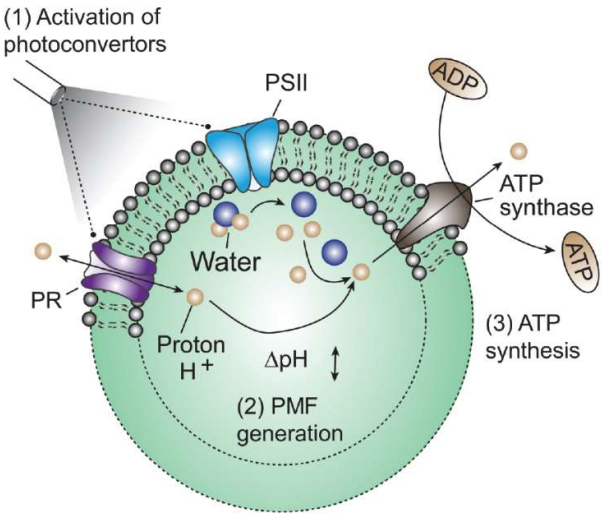

Figure 3. Artificial photosynthetic cells Schematic of artificial photosynthetic cell. Upon illumination, the vesicle synthesizes ATP by the coordinated activation of two complementary photoconverters (photosystem II, PSII and proteorhodopsin, PR) and an ATP synthase. PSII is activated by red light and acidifies the vesicle lumen, which allows the synthesis of ATP from ADP plus inorganic phosphate to take place on the outside. PR is activated by green light, which at low pH generates an electrochemical proton gradient, inside alkaline and negative, and thus impedes the synthesis of ATP. Figure taken with permission from [9].

(*Eq. 4*). Biochemical studies of decarboxylation reactions have shown two different mechanisms of energy conservation. In the first, the decarboxylation energy is converted directly into an electrochemical Na^+ gradient (Fig. 2A), as first shown for oxaloacetate decarboxylation by Peter Dimroth.[27] In the second mechanism, the substrate is decarboxylated and the substrate and product are exchanged across the membrane (Fig. 2B).[29,31] Since the substrate and product carry a different net charge (Table I), the antiport reaction generates a membrane potential. The chemistry of the decarboxylation reaction requires a proton, hence the formation of a pH gradient when the reaction is performed in confinement, *i.e.* inside a vesicle system. In a variation on this mechanism, it was demonstrated that monoanionic malate is taken up by uniport and the formed lactic acid leaves the vesicles by passive diffusion (Fig. 2B). In general, biological membranes are highly permeable for weak acids and passive fluxes are considerable, even when the ambient pH is 2-3 pH units higher than the pK_A of the relevant conjugate acid-base pair.[41] The energetics of the antiport and uniport is the same, but kinetically it can be advantageous to use an antiport mechanism as the product gradient contributes to the driving force for the influx of substrate and *vice versa*.

We have purified the malate/lactate antiporter and malolactic enzyme from *Lactococcus lactis* and reconstituted the system in synthetic lipid vesicles. A pH gradient and membrane potential are formed when the vesicles are supplied with L-malate. The co-reconstitution of the decarboxylation pathway together with the arginine breakdown pathway would represent two orthologous routes for metabolic energy conservation, allowing the synthetic cell to use both ATP and a proton motive force without the involvement of an ATP synthase/hydrolase. Table I shows that substrate/product antiport or exchange always involves a product that is more positively charged than the substrate, hence the $\Delta\Psi$ formed is inside negative relative to outside. The decarboxylation reaction inside the vesicles results in a ΔpH inside alkaline relative to outside. The arginine breakdown pathway can lead to acidification when citrulline is formed, but by combining the arginine breakdown pathway with the decarboxylation pathway it should be possible to better maintain a neutral to slightly alkaline internal pH.

2.3. Artificial photosynthetic cells

Numerous groups have co-reconstituted F_0F_1-ATP synthase with bacteriorhodopsin to control the synthesis of ATP by light. A disadvantage of this system is that the orientation of the proteins in the membrane is difficult to control. Recently, more advanced systems have been built with the aim of maintaining and controlling the electrochemical proton gradient. Shin and colleagues used the ATP synthase with two photoconverters, a photosystem II and proteorhodopsin.[9] The three proteins were reconstituted in small lipid vesicles ("artificial organelles") with the F_1 domain of the ATP synthase on the outside (Fig. 3). Upon activation of photosystem II by red light protons are pumped into the vesicles (the interior becomes positive and acidic), and the Δp drives the synthesis of ATP. Activation of proteorhodopsin by green light dissipates the Δp or even reverses the polarity of the electrochemical proton gradient, which impedes the synthesis of ATP. The artificial organelles were encapsulated in giant vesicles to provide them with ATP and drive endergonic reactions, such as pyruvate carboxylase-mediated carbon fixation and actin polymerization.

In another study, ATP synthase and bacteriorhodopsin were incorporated in small vesicles and used to drive protein synthesis in giant-unilamellar vesicles.[42] Remarkably, part of the *de novo* synthesized bacteriorhodopsin and ATP synthase were integrated into the artificial photosynthetic organelle and thereby enhanced the energetic capacity of the system. The proteins are synthesized by the components of the PURE system, but the machinery (Sec, YidC) for insertion of proteins into the membrane is missing. It remains to be established how the membrane proteins are (spontaneously) inserted in the artificial organelle membrane.

2.4. Molecular rheostat

In the arginine breakdown pathway, a remarkable degree of energy homeostasis is achieved, but the actual ATP level is influenced by the amount of ATP demanding reactions.[25] Bowie and colleagues have described a molecular rheostat that accounts for the ATP demand through switching between an ATP-generating and non-ATP-generating pathway according to the concentration of inorganic phosphate (Fig. 4, taken from [5]). The system is based on fourteen purified enzymes in a cell-free system and used to produce in solution isobutanol from glucose. The breakdown of glucose is branched at the level of glyceraldehyde-phosphate dehydrogenase (GAPdh) to make the use of NADH and ATP stoichiometrically balanced. In brief, in one branch the glyceraldehyde-3-phosphate (G3P) is metabolized via.GAPdh and phosphoglycerate kinase (PGK), yielding ATP and reducing equivalents. In the other branch G3P is converted via a non-phosphorylating glyceraldehyde dehydrogenase (GapN). GapN eliminates the production of ATP and generates NADPH rather than NADH, which is needed for the production of 2-ketoacid *iso*butanol. The relative flow through the ATP-generating branch is set by the concentration of inorganic phosphate, which is a substrate of GAPdh but not of GapN. Hence, the rheostat responds to the depletion of ATP and restores the ATP level by switching between the branches.

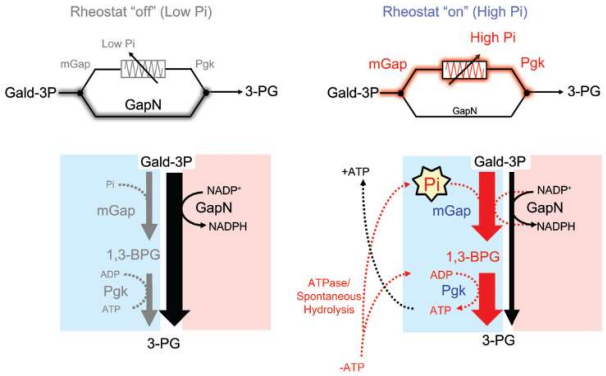

Figure 4. Molecular rheostat to control the ATP and NAD(P)H levels Schematic of the operation of the molecular rheostat. Left panel: at low Pi concentrations and high levels of ATP, the GapN pathway is used which generates no additional ATP. Right panel: at high Pi concentrations (resulting from the hydrolysis of ATP), the mGapDH– PGK pathway is used to restore the ATP level. G3P, glyceraldehyde-3 phosphate; 3PG, 3-phosphoglycerate; 1,3-BPG, 1,3-bisphosphoglycerate. Figure taken with permission from [5]

3. Compartmentalization and vesicle systems

3.1. Building blocks for membranes

One of the hallmarks of living species is compartmentalization, which implies that membrane-bounded systems may have arisen early on in the emergence life.[43] Compartmentalization in the form of vesicles allows molecules to concentrate, interact and coevolve, which is a *conditio sine qua non* for life. Vesicle structures can form spontaneously from fatty acids, as first reported in 1973,[44] and such membranes may have surrounded the first cells. Fatty acid-based vesicles are capable of growth and division when the appropriate components are added to the medium or the right physical conditions are imposed,[45,46] but they are less stable and more permeable to small molecules than conventional phospholipid-based membranes. Fatty acid-phospholipid blended membranes display increased stability but still maintain permeability for small (charged) solutes. They may have formed an intermediate in protocellular evolution, which allowed membrane passage without transporters.[47]

Well-sealed, stable membranes can also be formed from block copolymers,[48] but the functional incorporation of integral membrane proteins is challenging, especially when the proteins require specific lipids as cofactors. The majority of successful reconstitutions in non-native-amphiphile membranes involve relatively stable membrane pores or channels that do not undergo large conformational changes in the membrane.[49] Functional reconstitution of more complex enzyme systems has been achieved by using a blend of phospholipids and a block copolymer to stabilize the activity of the protein.[50] Today's biological membranes are mostly composed of lipids, in which proteins are embedded. Even if the reconstitution of complex membrane proteins in a block copolymer lipid blend would be possible, the synthesis (and incorporation) of block copolymers in a growing cell would require biochemical machinery that does not exist in organisms know today. Most vesicle systems for functional reconstitution use phospholipids.

We have studied numerous membrane transporters, both ATP- and electrochemical ion gradient-driven, and find that anionic lipids (phosphatidylglycerol or phosphatidylserine) and the non-bilayer lipid phosphatidylethanolamine are generally required for activity.[51] Many eukaryotic proteins require sterols for full functionality and cholesterol (mammalian), ergosterol (yeast) or plant-based sterols can be included in the reconstitution mixture.[52] For the hydrophobic chains we typically use 1,2-dioleoyl (diC18:1 Δ9-*cis*) or 1-palmitoyl-2-oleoyl (C16:0, C18:1 Δ9-*cis*), thus DOPX and POPX, respectively. DOPX membranes have a lower phase transition temperature and are less stable and more permeable for small molecules than POPX membranes.[41] Both at the level of lipid mixtures and at the level of blends between lipids and fatty acids or block copolymers, there is still a lot to be learned to enable (more) complex reconstitution of synthetic cell-like systems.

3.2. Membrane crowding

Biological membranes are highly crowded with proteins and thus the lipid-to-protein ratios are low; in the plane of the membrane only a few lipids separate individual protein complexes. For example, the weight-based lipid-to-protein ratio of the plasma membrane is about 1[53], which leaves about fifty lipids per leaflet to cover the perimeter of a 70kDa protein. Given that membrane proteins perturb the dynamics of lipids, a crowded biological membrane will be more rigid and less fluid than that of "dilute" liposomes, in which proteins are typically present at lipid-to-protein ratios of 10 to 1000 (w/w), corresponding to molar ratios of 1000 to 100,000. In synthetic vesicles with 2000 rather than 10,000 or more phospholipids per membrane protein (complex), the diffusion coefficient of lipids is already reduced by 20% and that of polytopic membrane proteins by 50%,[54] which is indicative of a lower fluidity or higher lipid order in the membrane. A lower fluidity may impact the (detergent-mediated) insertion of a protein into the membrane, which is the commonly used method of membrane reconstitution.[51,55] In fact, we find that the activity of membrane transport proteins does not increase proportionally with the amount of protein used for the reconstitution when the lipid-to-protein ratios fall below 2000 (mol/mol) .[56] This ratio corresponds to about 1500 proteins per μm^2 [54] and compares to 25,000 proteins per μm^2 in native plasma membranes. Apparently, not all proteins are correctly inserted into the membrane when the lipid-to-protein ratio drops below 2000.

Thus, our reconstitution technology may become a bottleneck in the bottom-up construction of synthetic cells when (multiple) proteins need to be incorporated at high concentrations. Ultimately, we will need protein insertion machineries like Sec[57] rather than detergent-destabilization of vesicles to build more complex systems.

3.3. Vesicle systems

Cell-sized aqueous compartments for synthetic cells range from submicrometer (large unilamellar vesicles, LUVs) to micrometer (giant-unilamellar vesicles, GUVs). Procedures have been developed to incorporate integral membrane proteins or lipid-anchored proteins into the membrane and to include enzymes and small molecules into the vesicle lumen. We typically form LUVs via detergent-mediated reconstitution ,[51] which is based on a method originally developed by Jean-Louis Rigaud.[55] We have produced sub-micron and micrometer size proteoliposomes with up to 50 mg/ml of protein or cell lysate in the vesicle lumen,[58] but technically it is challenging to achieve *in vivo*-like crowding levels (200-300 mg/mL).[59] By increasing the outside osmolality the vesicles shrink due to water efflux and

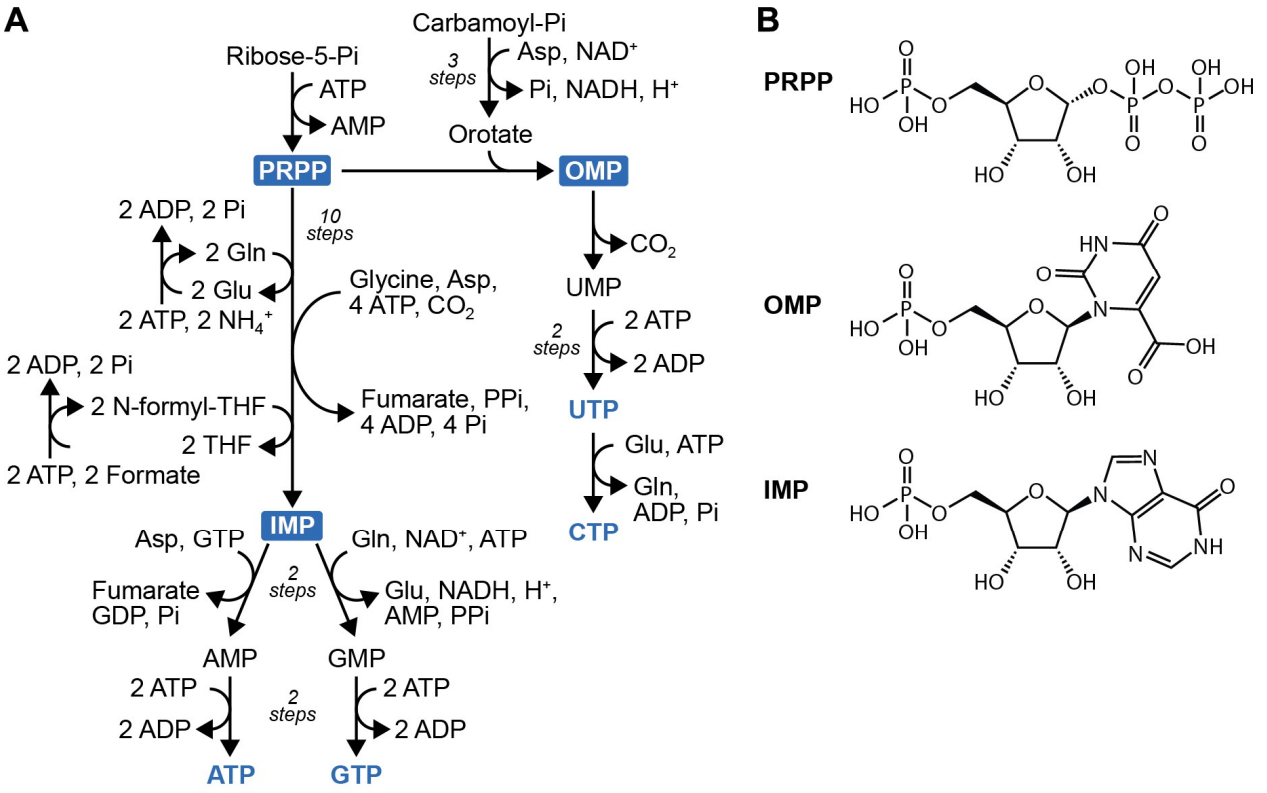

Figure 5. Building blocks for information carriers Synthesis of nucleotide triphosphates. (**A**) Simplified reaction diagram for the synthesis of ATP, GTP, UTP and CTP. After ribose-5-phosphate is converted into phosphoribosyl pyrophosphate (PRPP), it is converted in ten steps into inosine monophosphate (IMP) to form ATP and GTP. PRPP plus orotate yields orotidine 5'-monophosphate (OMP), which is converted in two steps into UTP and CTP. Gln, glutamine; Glu, glutamate; Asp, aspartate; THF, tetrahydrofolate. (**B**) Chemical structures of PRPP, OMP and IMP.

the luminal contents are concentrated. The shrinking of the vesicles is reversible, which occurs when osmolytes are taken up or the outside osmolality is reduced.[25] In this way one can study synthetic metabolic networks under varying conditions of crowding, ionic strength and osmotic pressure.

The sub-µm size lipid vesicles are robust and suitable for ensemble measurements of solute import, cargo release, and single-liposome analysis of vesicle size and swelling,[60] and recently LUVs have been used to reconstitute a metabolic network for energy and physicochemical homeostasis.[25] Although LUVs are small, they have dimensions similar to that of small, free-living bacteria such as *Pelagibacter,* and thus their volume should not pose a hurdle for accommodating all the essential components of a cell (see Section 4). The µm-size GUVs are more fragile but offer the advantage that they can be used for patch clamp and light microscopy studies. Membrane domain formation, the dynamics of individual molecules and their possible interaction with other membrane components can be tracked.[61] In the context of bottom-up synthetic biology GUVs have been used as platform to develop artificial photosynthetic organelles,[9] synthetic beta-cells[62] and motile light-guided synthetic cells.[63]

3.4. A metabolic network for energy and physicochemical homeostasis

Any living cell maintains the pH, ionic strength, osmotic pressure, macromolecular crowding and ΔG_p within limits to allow the enzymes and other components to function near their optimum. Hence, the importance to obtain physicochemical homeostasis in cell-like systems. The arginine breakdown pathway has been co-reconstituted with an ionic strength-gated ATP-driven osmolyte transporter to allow vesicle expansion and restoration of the physical chemical conditions upon exposure to osmotic stress.[25] When the vesicles are exposed to an increasing medium osmolality, they shrink and the ionic strength increases and the concentrations of the internal components are increased. Under these conditions the pathway functions suboptimally and the enzymes are gradually inactivated. However, when the ionic strength reaches a critical value, the ATP-driven osmolyte transporter is activated and glycine betaine is pumped inside, which is accompanied by passive influx of water into the vesicles. This increases the volume, reduces the ionic strength and stabilizes the internal pH and thus enables basic physicochemical homeostasis.

4. How much ATP does a synthetic cell need?

One of the essential design factors of synthetic cells is the amount of energy required for the cell to perform its (core) functions. As an example, in *E. coli* the ATP turnover is a few million molecules per second, given that the ATP pool is turned over 4 to 7.5 times per second[64], a volume of 1 fL[65] and an internal ATP concentration of 10mM[65]. In section 4, we elaborate on the energy requirements of a hypothetical synthetic cell, focusing on the quantification of the ATP-consuming reactions. First, we list important energy requiring processes, and in the second part we make a quantification of the ATP equivalents needed to operate a synthetic cell. We list all energy used by a cell in terms of ATP equivalents (Table 2) as it takes one ATP to regenerate GDP (or any other nucleotide-diphosphate) to the triphosphate form by a nucleoside-diphosphate kinase. We estimate that of all nucleotides turned over about 80% is in the form of ATP.

4.1. Proteins

In bacteria, the vast majority of all ATP (around 75%) is used for the synthesis of proteins.[66] Most of that energy is used for the synthesis of ribosomes and formation of the peptide bond. The energy that is used for synthesis of amino acids, can be minimized by taking up amino acids in the form of di- or tripeptides, followed by internal digestion through peptidases. The membrane transporter DtpT takes up virtually every di- or tripeptide together with one or multiple protons, driven by Δp.[67] An alternative broad specificity transporter Opp, belonging to the ABC superfamily, imports oligopeptides with lengths between 4-35 amino acids,[68] likely using 2 ATP equivalents per oligopeptide. Digestion of these di, tri or oligopeptides into amino acids can then be done by amino- and endopeptidases, without additional energy cost. This lowers the metabolic energy cost for synthesis of amino acids to less than 1 ATP per amino acid. Forming a new peptide bond however requires approximately 4 ATP equivalents. Two ATP equivalents for amino acid activation, one ATP equivalent for aminoacyl-tRNA binding to the elongation factor and finally one ATP equivalent for the translocation reaction, where the peptidyl-tRNA is translocated from the A-site to the P-site.[69,70]

4.2. Information carriers

Nucleotides for information carriers can be synthesized *de novo*, using approximately 50 ATP equivalents per nucleotide.[71] The energy costs are lower when a simpler route is used (Fig. 5) and the necessary amino acids are imported (section 3.1). The energy cost of the simplified pathway is around 10 ATP equivalents per nucleotide (Fig. 5). Here, the conversion of ribose-5-Pi, carbamoyl phosphate and amino acids to the final products (ATP, GTP UTP and CTP) requires around 20 enzymatic steps, which is manageable from an engineering perspective. The main drawback of *de novo* nucleotide synthesis is that it comes with the complex regulation of pathways and the underlying biochemistry of different components.

An even simpler solution than outlined in Figure 5 is to take up the nucleotides directly from the environment, a solution used by pathogens that lost their ability to synthesize their nucleotides.[72] By for example using a combination of the nucleotide carriers PamNTT3 and PamNTT5, the nucleotides UTP, GTP and ATP can be taken up with Δp as a driving force.[73] UTP can then be converted into CTP, using one ATP equivalent in a single enzymatic step. After the nucleotide-triphosphates are converted by nucleoside-diphosphate kinase into nucleotide-diphosphates, they can then be turned into their respective deoxyribonucleotides analogues. Using this strategy, the nucleotides can be produced by a minimal set of enzymes requiring less than 3 ATP equivalents per nucleotide. dUMP can be converted into dTMP by a thymidylate synthetase, after which dTMP is converted into dTTP. Apart from the metabolic energy cost for synthesis or and import of nucleotides, the formation and maintenance of DNA and RNA have additional energy costs. For DNA, the error correction is estimated to require one ATP equivalent per built-in nucleotide.[71] For mRNA, the degradation rate needs to be considered, as the lifetime of for instance mRNA is shorter than the cell cycle. When mRNA is degraded, the nucleotides can be recycled, which takes 2 ATP equivalents per nucleotide.[71]

4.3. Lipid synthesis for compartmentalization

The minimal lipid composition of a synthetic cell consists of 50% DOPE (*1,2-dioleoyl-sn-glycero-3-phosphoethanolamine*) plus 50% DOPG (*1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)*). This lipid composition supports high rates of transport of the bacterial transporters that we have studied; for eukaryotic membrane proteins a sterol and some specific lipids may be required (see section *4.1*), which we do not consider here. Synthesis of these lipids, or similar ones with different acyl chains (*e.g.* POPE and POPG), can be performed by combining a set of around ten enzymes.[74] Starting from oleic acid and glycerol, the intermediate CDP-DAG is formed in four enzymatic steps, after which further conversion yields either DOPE or DOPG (Fig. 6). Coenzyme A is required for the lipid synthesis but is also regenerated by FadD. The initial amount of coenzyme A can be synthesized *de novo* from pantothenate, or imported using an acetyl-CoA transporter, *e.g.* ACATN1.[75] In total the synthesis of DOPE and DOPG by this pathway takes 7 and 8 ATP equivalents per lipid, respectively. Adding a lipid scramblase would enforce the lipids to distribute over both the inner and outer leaflets.[76]

4.4. Membrane transport for osmotic, ionic and pH control

Growing (synthetic) cells should maintain their osmolarity, ionic strength, and pH in order to keep the cellular machinery active and maintain a stable, out-of-equilibrium state. Therefore, import of ions, compatible solutes and inorganic phosphate (P_i) is crucial. The most abundant ions in cells are K^+ (30-300 mM) and Mg^{2+} (30-100 mM) as cations, and inorganic and organic phosphates (~100 mM), glutamate (100mM), RNA, DNA and proteins as anions.[65] Except for RNA, DNA and proteins these ions need to be taken up by membrane transporters, mostly driven by Δp *e.g.* the phosphate transporters of the PiT family[77]; ATP *e.g.* the high affinity potassium uptake system Kdp[78]; or both *e.g.* the Trk potassium uptake system.[79] Here, for simplicity we count one ATP per ion that is taken up.

4.5. Maintenance energy

Maintenance costs cover the energy that is spent on anything that is not directly related to growth. For example: energy loss in futile cycling of enzymes, leakage of compounds over the membrane or processes like adaptation, *e.g.* pH and osmoregulation to keep the cytosolic conditions right. The maintenance energy can be estimated from the energy

Box I. Sensors to measure the energy and physicochemical status of cells

Several genetically encoded sensors and chemical probes are available to monitor the energy and redox status and physicochemical conditions of synthetic cells. Here, we describe some generic sensors used in our synthetic biochemistry program; numerous solute-specific sensors are described in references [80,81]

ATP: The ATeam sensors are FRET based and consist of two fluorescent proteins (FPs), which are connected by the ε-subunit of the F_0F_1-ATP synthase from *Bacillus subtilis*.[82] Upon binding of ATP the ε-subunit adopts a compact conformation and draws the two fluorophores closer together, increasing the FRET ratio. Three variants are available with high and low affinity for ATP, and a version that does not bind ATP. A single fluorophore variant has been developed in which the readout is provided by a single circular permutated FP.[83] PercevalHR binds ATP and ADP with similar, micromolar affinities. At physiological levels of adenine nucleotides PercevalHR is practically fully saturated with ligand and therefore reports the ATP to ADP ratio rather than the absolute concentration of ATP or ADP.[84] As in the Queen sensors, a circular permutated FP allows ratiometric readout. Lastly, based on Queen, an intensiometric ATP sensor was developed which can be bound to the membrane.[85]

NAD(P)H: SoNar is a ratiometric genetically encoded sensor that reports the NAD^+ to NADH ratio.[86] iNap is a derivative of SoNar and reports the NADPH concentration instead of the ratio between $NADP^+$ and NADPH.[87]

pH: pHluorin and pHred are protein-based pH sensors.[88,89] pHluorin is based on GFP and has spectral properties in the yellow and green region. pHred is based on mKeima and is, owing to its large stokes shift, compatible with the ATP sensor PercevalHR. In addition to protein-based sensors, chemical probes are available like pyranine and BCECF.[90,91] These are commercially available and allow imaging for longer periods of time than the protein-based sensors. Methyl-ester derivatives of BCECF readily permeate the plasma membrane, and in the cytosol the molecules become trapped upon hydrolysis of the ester bond (esterase activity). Given the value of the external pH, measurements of the internal pH enable calculation of the magnitude of the ΔpH across the membrane.

Membrane potential: The membrane potential (ΔΨ) is measured by chemical probes, like $diSC_3$-5.[92] The exact mechanism of how $diSC_3$-5 reports changes in the ΔΨ is not fully understood, but its fluorescent intensity increases upon interaction with lipid membranes. This fluorescence is quenched upon polarization of the membrane. The magnitude of the proton motive force is obtained by combining the ΔΨ & ΔpH, according to equation 2.

Ionic strength: The ionic strength is measured with a FRET sensor that consists of two fluorescent proteins joined by a flexible linker and two α-helices with opposite charges.[93] The FRET signal is high when the ionic strength is low, and the signal is low when a high ionic strength of the solution shields the charges of the α-helices.

Excluded volume: The excluded volume or so-called macromolecular crowding sensors have a similar design as the ionic strength sensor, except that the same charge pairs are present on both α-helices. Here, the excluded volume drives a more compact state of the sensor, which is observed as an increase in FRET signal.[94,95] A similar crowding-sensing principle was used in a synthetic sensor; here, two chemical fluorophores forming a FRET pair are connected by a polyethylene polymer linker.[96]

Viscosity: Viscosity is measured by fluorescent molecular rotors. These rely on intra-molecular rotation, which is suppressed by a high viscosity, which results in increased fluorescence. Fluorescent molecular rotors are available as intensio- and ratiometric sensors.[97,98]

Potassium: KIRIN1/KIRIN-GR and GINKO1 are potassium ion sensors that are based on the same K^+ binding protein, but differ in the fluorescent proteins used. The two KIRIN sensors use different FRET pairs, whereas GINKO1 has only one circular permutated FP.[99] They report potassium concentrations in the low millimolar range.

Parameter	*Name*	*Fluorophore*	*Read-out*	*Spectral maxima (nm)*		*Comments*
				excitation	*emission*	
ATP	Ateam	CFP mVenus	Ratiometric FRET	435	475 527	Moderately pH sensitive
ATP	Queen	cpEGFP	Ratiometric	400 494	513	Moderately pH sensitive
ATP/ADP	PercevalHR	cpmVenus	Ratiometric	420 500	515	pH sensitive
ATP	iATPSnFR	cpSFGFP	Intensiometric	490	512	Ratiometric when fused to mRuby, moderately pH sensitive
NAD+/NADH	SoNar	cpYFP	Ratiometric	420 485	530	pH insensitive
NADPH	iNAP	cpYFP	Ratiometric	420 485	530	pH insensitive
pH	pHluorin	GFP	Ratiometric	410 470	535	Intensiometric variant available
pH	pHred	mKeima	Ratiometric	440 585	610	Compatible with PercevalHR
pH	pyranine	Arylsulfonate	Ratiometric	400 450	510	Commercially available
pH	BCECF	fluorescein	Ratiometric	439 490	530	Commercially available
Ionic Strength	I-sensor	Cerulean Citrine	Ratiometric FRET	420	475 525	Different designs available
Excluded volume	Crowding sensor	Cerulean Citrine	Ratiometric FRET	420	475 525	Sensors differing in crowding sensitivity are available; different designs available
Excluded volume	Synthetic crowding sensor	Atto488 Atto565	Ratiometric FRET	470 555	512 630	Not commercially available
Membrane Potential	$DiSC_3$-5	carbocyanine	Intensiometric	653	676	Commercially available
Viscosity	Various					Various classes available, including ratiometric variants
K+	KIRIN1	mCerulean3 cpVenus	Ratiometric FRET	410	475 530	Selective for K^+ over Na^+
K+	KIRIN-GR	Clover mRuby2	Ratiometric FRET	470	520 600	Small FRET change
K+	GINKO1	EGFP	Ratiometric	400 500	520	Sensitive to high concentrations of Na^+

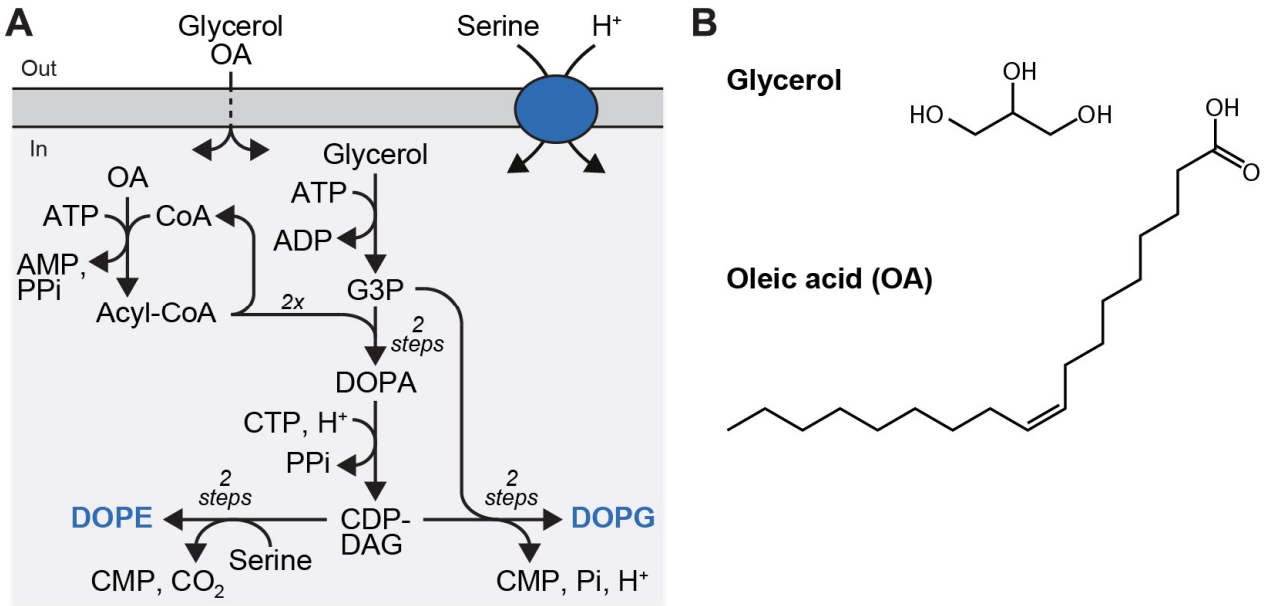

Figure 6. Lipid biosynthesis Synthesis of two major phospholipids, DOPG and DOPE. (**A**) Reaction diagram for the synthesis of *1,2-dioleoyl-sn-glycero-3-phosphoethanolamine* (DOPG) and *1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)* (DOPG) from the precursors glycerol, oleic acid and serine. Both glycerol and oleic acid (OA) can diffuse across the membrane, after which they are converted into glycerol 3-phosphate (G3P) and acyl-coenzyme A (acyl-CoA), respectively. Two molecules of acyl-CoA react with G3P to form *1,2-dioleoyl-sn-glycero-3-phosphate* (DOPA), from which DOPE and DOPG can be formed in three steps. CDP-DAG, cytidine diphosphate diacylglycerol. Phospholipids with alternative acyl chains can be synthesized by feeding the synthetic cell with the appropriate fatty acids. (**B**) Chemical structures of glycerol and oleic acid (OA).

(uptake) at various growth rates by extrapolation to zero growth. Measurement or quantification of this parameter is not straightforward, since it varies depending on the specific metabolism. Feist et al. report, based on a metabolic reconstruction of aerobically growing *E. coli* cells, a non-growth associated maintenance (NGAM) of 8.4 ATP/gDW/h, while the growth associated energy costs are 59.8 ATP/gDW/h.[100] If we take the weight of one *E.* coli to be 1 *p*g[65] the ATP consumption for NGAM of a single cell is $4.8 \cdot 10^9$ ATP equivalents per hour compared to $3.4 \cdot 10^{10}$ ATP equivalents for the growth-associated costs.

4.6. Quantification of ATP demand of minimal synthetic cell

To quantify the energy requirement of a synthetic cell, we assume a spherical cell with a diameter of 400 nm and a volume of 0.03 *f*L, a size comparable to the size of the smallest free-living micro-organisms known today.[101] To estimate the protein content, we assume that the crowding is comparable to that of *e.g. E. coli*, which has approximately $3 \cdot 10^6$ proteins per μm^3 and a volume of about 1 *f*L.[65] A synthetic cell with a volume of 0.03 fl would thus contain 10^5 proteins. The average protein has a length of 300 amino acids and costs 5 ATP equivalents per amino acid. If we assume that the lifetime of a protein is longer than the cell cycle then the synthesis of all proteins takes $1.5 \cdot 10^8$ ATP equivalents.

We quantify the DNA replication and transcription by taking a genome size of 500 genes, similar to the genome of JCVI-syn 3.0.[3] We take an average gene length of 900 base pairs (300 amino acid protein and minimal intergenic DNA) and thus the genome would consist of $4.5 \cdot 10^5$ base pairs. Taking 3 ATP equivalents per nucleotide, the total energy cost for the genome would be $3.6 \cdot 10^6$ molecules of ATP.

The cost of transcription depends on the total RNA level, which for *E. coli* can be estimated at 10^3-10^4 copies per cell.[65] Following the same calculation, we estimate the synthetic cell to have 20-200 copies per cell, based on the aforementioned protein concentration and a doubling time of an hour. Since this would mean less than one transcript per protein we take a number of 500 copies per cell, which is equal to the protein number. If we take a degradation rate of 10^{-3} copies s^{-1} the total ATP consumption would be $4.6 \cdot 10^6$ ATP equivalents.[71] For rRNA and tRNA we take $1.0 \cdot 10^3$ and $1.3 \cdot 10^4$ copies[65] leading to an energy cost of $1.4 \cdot 10^7$ and $3.4 \cdot 10^6$ ATP equivalents, respectively.

The synthetic cell, spherical with a diameter of 400nm, has a surface area of $5 \cdot 10^5$ nm^2 requiring a bilayer of $1.5 \cdot 10^6$ phospholipids, assuming an area per lipid of 0.7 nm^2. The lipid synthesis starting from fatty acids, takes 7.5 ATP equivalents per lipid if we take a composition of 50% DOPE plus 50% DOPG. Thus, it would take $1.1 \cdot 10^7$ ATP equivalents to synthesize all lipids.

ATP required for transport and maintenance can be estimated as follows. If we sum all ions and solutes needed for maintaining the internal osmotic pressure, ionic strength and metabolite pool (260-600mM) and assume that uptake of each molecule costs one ATP, then a cell with a volume of 0.03 *f*L would consume 0.5 – $1.1 \cdot 10^7$ ATP equivalents. A single *E. coli* uses $4.8 \cdot 10^9$ ATP per hour for non-growth associated maintenance (see above). Assuming that this scales with cell volume, the synthetic cell requires 30 times less: $1.6 \cdot 10^8$ ATP equivalents

We think that the amount of ATP required for cell division is small compared to that of the other processes considered. The absolute amount of ATP is difficult to estimate, FtsZ being a dynamic system.[102] Taking estimations from Dr. DJ Scheffers (personal communication), 20.000 FtsZ per *E. coli* with a turnover of 1.5 GTP/min gives $7.5 \cdot 10^5$ GTP (or ATP-equivalents) per doubling (25 minutes.). For our synthetic cell with a doubling time of an hour and a volume that is 30 times smaller we estimate $6.0 \cdot 10^4$ ATP equivalents.

In Table II we compare the categorized ATP consumption of our hypothetical synthetic cell with that of *E. coli*.[103] We find as expected that a major fraction of the ATP is needed for the synthesis of protein, and surprisingly a similar amount of ATP is used for maintenance. In the synthetic cell the ATP needed for maintenance is based maintenance in *E. coli*, which is a much more complex organism than the synthetic cell, therefore scaling only to volume might well lead to an overestimation.

Table 2. ATP requirements of the major cellular processes. The data for *E. coli* in mmol/gram dry weight were taken from [103] and converted into ATP equivalents per cell, asuming a cytoplasmic volume of 1 *f*L. The synthetic cell data are based on a spherical cell-like system with a volume 0.03 *f*L. The quantification of the ATP costs for this system is described in section 4.6.

Process	*Synthetic cell (ATP equivalents)*	*E. coli (mmol/g dry wt)*	*E. coli (ATP equivalents)*
Protein			
- Uptake of amino acids	$3.0 \cdot 10^7$ (8%)		
- Glucose to amino acids		1.4	$8.0 \cdot 10^8$
- Translation	$1.2 \cdot 10^8$ (34%)	19.1	$1.1 \cdot 10^{10}$
DNA	$3.6 \cdot 10^6$ (1%)	1.1	$6.3 \cdot 10^8$
RNA		4.4	$3.3 \cdot 10^9$
mRNA	$4.6 \cdot 10^6$ (1%)		
tRNA	$3.4 \cdot 10^6$ (1%)		
rRNA	$1.4 \cdot 10^7$ (4%)		
Lipids	$1.1 \cdot 10^7$ (3%)	0.1	$5.7 \cdot 10^7$
Transport (other than amino acids)	$1.1 \cdot 10^7$ (3%)	5.2	$3.0 \cdot 10^9$
Maintenance	$1.6 \cdot 10^8$ (45%)		$4.8 \cdot 10^9$ *
Division	$6.0 \cdot 10^4$ (0%)		$7.5 \cdot 10^5$ **

* Maintenance from [100] ** Estimation by Dr. DJ Scheffers (see text).

5. Outlook and Perspectives

The construction of a living cell from molecular components is one of the major challenges of today's chemistry and life sciences, as one is crossing the border from the 'dead' molecules of chemistry to the living systems of biology. It has not yet been possible to rationally design and construct, using a bottom-up constructive approach, a simple form of life based on a limited number of molecular building blocks. While our fundamental understanding of the individual building blocks of life is rapidly growing, putting a minimal set of components together such that life-like properties emerge remains a formidable, yet exciting challenge.

Non-equilibrium systems are driven by the continuous flow of energy and matter and can develop into a multitude of states, *e.g.* when the flow of matter is perturbed. Nature is an assemblage of many of such open systems, each of which can take its own path. The challenge is to construct and control such systems. In this paper, we have presented an overview of the simplest systems one could envisage to sustainably supply a cell with fuel in the form of ATP and/or electrochemical ion gradients. By coupling the energy feed to product export, it is possible to maintain a continuous flow of in the pathways for ATP or ion gradient formation. One of the bottlenecks in current systems is that one or a few components, for instance ATP, runs out, leading the system to equilibrium. We have recently shown that it is possible to use the provision and consumption of ATP for physicochemical homeostasis in synthetic vesicles. The next challenge is to couple the metabolic energy conservation to synthetic modules for *e.g.* lipid, protein, and nucleic acid synthesis (Fig. 5), yet maintain energy and physicochemical homeostasis. Ultimately, the synthesis of the components needs to be directed by a synthetic genome, and we need to coordinate DNA replication with growth and division. In Box II we present a series of outstanding questions on fuel supply and homeostasis of metabolic energy in synthetic cells.

Box II. Open questions

1. Is the interconversion of ATP and electrochemical ion via ATPsynthase /hydrolase essential for life?
2. How much ATP is required for polymer synthesis and maintenance processes in small cell-like systems?
3. What is the lower limit in size for a cell?
4. What are the physicochemical limits for life of *e.g.* ionic strength or macromolecular crowding?
5. How can we increase the efficiency of membrane reconstitution and molecule encapsulation to build more complex cell-like systems?
6. Bridging the gap between bottom up and top down. What do we know?
7. How many unknown components are there still to be discovered?
8. How can we use bio-orthogonal systems in living systems?

Keywords: bottom-up construction • synthetic cell • metabolic energy conservation • cellular homeostasis • synthetic biochemistry

Acknowledgements

The work was funded by an ERC Advanced Grant (ABCvolume; #670578) and the Netherlands Organization for Scientific Research Gravitation program BaSyC.

[1] N. Lane, W. Martin, *Nature* **2010**, 467, 929-934.
[2] D. G. Gibson, J. I. Glass, C. Lartigue, V. N. Noskov, R.-Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z.-Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison, H. O. Smith, J. C. Venter, *Science* **2010**, 329, 52-56.
[3] C. A. Hutchison, R.-Y. Chuang, V. N. Noskov, N. Assad-Garcia, T. J. Deerinck, M. H. Ellisman, J. Gill, K. Kannan, B. J. Karas, L. Ma, J. F. Pelletier, Z.-Q. Qi, R. A. Richter, E. A. Strychalski, L. Sun, Y. Suzuki, B. Tsvetanova, K. S. Wise, H. O. Smith, J. I. Glass, C. Merryman, D. G. Gibson, J. C. Venter, *Science* **2016**, 351, aad6253.
[4] M. Porcar, A. Danchin, V. de Lorenzo, V. A. Dos Santos, N. Krasnogor, S. Rasmussen, A. Moya, *Syst. Synth. Biol.* **2011**, 5, 1-9.
[5] P. H. Opgenorth, T. P. Korman, L. Iancu, J. U. Bowie, *Nat. Chem. Biol.* **2017**, 13, 938-942.
[6] A. Bhattacharya, R. J. Brea, H. Niederholtmeyer, N. K. Devaraj, *Nat. Commun.* **2019**, 10, 300.
[7] V. Noireaux, Y. T. Maeda, A. Libchaber, *Proc. Natl. Acad. Sci. U. S. A.* **2011**, 108, 3473-3480.
[8] F. Caschera, V. Noireaux, *Curr. Opin. Chem. Biol.* **2014**, 22, 85-91.
[9] K. Y. Lee, S.-J. Park, K. A. Lee, S.-H. Kim, H. Kim, Y. Meroz, L. Mahadevan, K.-H. Jung, T. K. Ahn, K. K. Parker, K. Shin, *Nat. Biotechnol.* **2018**, 36, 530-535.
[10] M. Loose, T. J. Mitchison, *Nat. Cell Biol.* **2014**, 16, 38-46.
[11] T. Matsuura, K. Hosoda, Y. Kazuta, N. Ichihashi, H. Suzuki, T. Yomo, *ACS Synth. Biol.* **2012**, 1, 431-437.
[12] H. Ledford, *Nature* **2016**, 536, 136-137.
[13] P. Mohanraju, K. S. Makarova, B. Zetsche, F. Zhang, E. V Koonin, J. van der Oost, *Science* **2016**, 353, aad5147.
[14] N. A. Moran, G. M. Bennett, *Annu. Rev. Microbiol.* **2014**, 68, 195-215.
[15] V. Pérez-Brocal, R. Gil, S. Ramos, A. Lamelas, M. Postigo, J. M. Michelena, F. J. Silva, A. Moya, A. Latorre, *Science* **2006**, 314, 312-313.
[16] S. J. Ferguson, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 16755-16756.
[17] I. N. Watt, M. G. Montgomery, M. J. Runswick, A. G. W. Leslie, J. E. Walker, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, 107, 16823-16827.
[18] D. Pogoryelov, J. Yu, T. Meier, J. Vonck, P. Dimroth, D. J. Muller, *EMBO Rep.* **2005**, 6, 1040-1044.
[19] W. Junge, N. Nelson, *Annu. Rev. Biochem.* **2015**, 84, 631-657.
[20] C. Gibbons, M. G. Montgomery, A. G. Leslie, J. E. Walker, *Nat. Struct. Biol.* **2000**, 7, 1055-1061.
[21] A. J. Rodgers, M. C. Wilce, *Nat. Struct. Biol.* **2000**, 7, 1051-1054.
[22] S. P. Tsunoda, A. J. Rodgers, R. Aggeler, M. C. Wilce, M. Yoshida, R. A. Capaldi, *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 6560-6564.
[23] A. Meyrat, C. von Ballmoos, *Sci. Rep.* **2019**, 9, 3070.
[24] A. J. Driessen, B. Poolman, R. Kiewiet, W. Konings, *Proc. Natl. Acad. Sci. U. S. A.* **1987**, 84, 6093-6097.
[25] T. Pols, H. R. Sikkema, B. F. Gaastra, J. Frallicciardi, W. M. Śmigiel, S. Singh, B. Poolman, *manuscript submitted for publication* **2019**, DOI: 10.1101/698498.
[26] M. Berg, H. Hilbi, P. Dimroth, *Eur. J. Biochem.* **1997**, 245, 103-115.
[27] P. Dimroth, *Eur. J. Biochem.* **1982**, 121, 443-449.
[28] W. Hilpert, B. Schink, P. Dimroth, *EMBO J.* **1984**, 3, 1665-1670.
[29] V. Anantharam, M. J. Allison, P. C. Maloney, *J. Biol. Chem.* **1989**, 264, 7244-7250.
[30] T. Hirai, J. A. W. Heymann, D. Shi, R. Sarker, P. C. Maloney, S. Subramaniam, *Nat. Struct. Biol.* **2002**, 9, 597-600.
[31] B. Poolman, D. Molenaar, E. J. Smid, T. Ubbink, T. Abee, P. P. Renault, W. N. Konings, *J. Bacteriol.* **1991**, 173, 6030-6037.
[32] M. Salema, B. Poolman, J. S. Lolkema, M. C. L. Dias, W. N. Konings, *Eur. J. Biochem.* **1994**, 225, 289-295.
[33] M. Salema, I. Capucho, B. Poolman, M. V San Romão, M. C. Dias, *J. Bacteriol.* **1996**, 178, 5537-5539.
[34] H. Ilgü, J.-M. Jeckelmann, V. Gapsys, Z. Ucurum, B. L. de Groot, D. Fotiadis, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, 113, 10358-10363.
[35] R. L. Chisholm, R. a Firtel, *Nat. Rev. Mol. Cell Biol.* **2004**, 5, 531-541.
[36] D. Ma, P. Lu, C. Yan, C. Fan, P. Yin, J. Wang, Y. Shi, *Nature* **2012**, 483, 632-636.
[37] D. Molenaar, J. S. Bosscher, B. ten Brink, A. J. Driessen, W. N. Konings, *J. Bacteriol.* **1993**, 175, 2864-2870.
[38] A. Romano, H. Trip, J. S. Lolkema, P. M. Lucas, *J. Bacteriol.* **2013**, 195, 1249-1254.
[39] M. Coton, M. Fernández, H. Trip, V. Ladero, N. L. Mulder, J. S. Lolkema, M. A. Alvarez, E. Coton, *Microbiology* **2011**, 157, 1841-1849.
[40] P. Dimroth, B. Schink, *Arch. Microbiol.* **1998**, 170, 69-77.
[41] M. Gabba, J. Frallicciardi, J. S. van 't Klooster, R. Henderson, Ł. Syga, R. Mans, A. J. A. Maris, B. Poolman, *manuscript submitted for publication* **2019**.
[42] S. Berhanu, T. Ueda, Y. Kuruma, *Nat. Commun.* **2019**, 10, 1325.
[43] J. Spitzer, B. Poolman, *Microbiol. Mol. Biol. Rev.* **2009**, 73, 371-388.
[44] J. M. Gebicki, M. Hicks, *Nature* **1973**, 243, 232-234.
[45] I. Budin, A. Debnath, J. W. Szostak, *J. Am. Chem. Soc.* **2012**, 134, 20812-20819.
[46] C. Hentrich, J. W. Szostak, *Langmuir* **2014**, 30, 14916-14925.
[47] L. Jin, N. P. Kamat, S. Jena, J. W. Szostak, *Small* **2018**, 14, e1704077.
[48] A. Puiggalí-Jou, L. J. Del Valle, C. Alemán, *Soft Matter* **2019**, 15, 2722-2736.
[49] P. A. Beales, S. Khan, S. P. Muench, L. J. C. Jeuken, *Biochem. Soc. Trans.* **2017**, 45, 15-26.
[50] R. Seneviratne, S. Khan, E. Moscrop, M. Rappolt, S. P. Muench, L. J. C. Jeuken, P. A. Beales, *Methods* **2018**, 147, 142-149.
[51] E. R. Geertsma, N. A. B. Nik Mahmood, G. K. Schuurman-Wolters, B. Poolman, *Nat. Protoc.* **2008**, 3, 256-266.
[52] F. Bianchi, J. S. van 't Klooster, S. J. Ruiz, K. Luck, T. Pols, I. L. Urbatsch, B. Poolman, *Sci. Rep.* **2016**, 6, 31443.
[53] A. D. Dupuy, D. M. Engelman, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, 105, 2848-2852.
[54] S. Ramadurai, A. Holt, V. Krasnikov, G. van den Bogaart, J. A. Killian, B. Poolman, *J. Am. Chem. Soc.* **2009**, 131, 12650-12656.
[55] J. L. Rigaud, B. Pitard, D. Levy, *Biochim. Biophys. Acta* **1995**, 1231, 223-246.
[56] J. Knol, K. Sjollema, B. Poolman, *Biochemistry* **1998**, 37, 16410-16415.
[57] A.-B. Seinen, A. J. M. Driessen, *Annu. Rev. Biophys.* **2019**, 48, 185-207.
[58] G. van den Bogaart, N. Hermans, V. Krasnikov, B. Poolman, *Mol. Microbiol.* **2007**, 64, 858-871.
[59] J. van den Berg, A. J. Boersma, B. Poolman, *Nat. Rev. Microbiol.* **March 2017**.
[60] G. van den Bogaart, I. Kusters, J. Velásquez, J. T. Mika, V. Krasnikov, a. J. M. Driessen, B. Poolman, *Methods* **2008**, 46, 123-130.
[61] M. K. Doeven, J. H. A. Folgering, V. Krasnikov, E. R. Geertsma, G. van den Bogaart, B. Poolman, *Biophys. J.* **2005**, 88, 1134-1142.
[62] Z. Chen, J. Wang, W. Sun, E. Archibong, A. R. Kahkoska, X. Zhang, Y. Lu, F. S. Ligler, J. B. Buse, Z. Gu, *Nat. Chem. Biol.* **2018**, 14, 86-93.
[63] S. M. Bartelt, J. Steinkühler, R. Dimova, S. V. Wegner, *Nano Lett.* **2018**, 18, 7268-7274.
[64] W. H. Holms, I. D. Hamilton, A. G. Robertson, *Arch. Mikrobiol.* **1972**, 83, 95-109.
[65] R. Milo, R. Phillips, *Cell Biology by the Numbers*, Garland Science, **2016**.
[66] F. M. Harold, *The Vital Force: A Study of Bioenergetics*, **1986**.
[67] G. Fang, W. N. Konings, B. Poolman, *J. Bacteriol.* **2000**, 182, 2530-2535.
[68] F. J. Detmers, F. C. Lanfermeijer, R. Abele, R. W. Jack, R. Tampe, W. N. Konings, B. Poolman, *Proc. Natl. Acad. Sci. U. S. A.* **2000**, 97, 12487-12492.
[69] K. A. Calhoun, J. R. Swartz, *Methods Mol. Biol.* **2007**, 375, 3-17.
[70] H.-C. Kim, D.-M. Kim, *J. Biosci. Bioeng.* **2009**, 108, 1-4.
[71] M. Lynch, G. K. Marinov, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 112, 15690-15695.
[72] M. Horn, A. Collingro, S. Schmitz-Esser, C. L. Beier, U. Purkhold, B. Fartmann, P. Brandt, G. J. Nyakatura, M. Droege, D. Frishman, T. Rattei, H.-W. Mewes, M. Wagner, *Science* **2004**, 304, 728-730.
[73] I. Haferkamp, S. Schmitz-Esser, M. Wagner, N. Neigel, M. Horn, H. E. Neuhaus, *Mol. Microbiol.* **2006**, 60, 1534-1545.
[74] M. Exterkate, A. Caforio, M. C. A. Stuart, A. J. M. Driessen, *ACS Synth. Biol.* **2018**, 7, 153-165.
[75] Y. Hirabayashi, K. H. Nomura, K. Nomura, *Mol. Aspects Med.* **2013**, 34, 586-589.
[76] C. Alvadia, N. K. Lim, V. Clerico Mosina, G. T. Oostergetel, R. Dutzler, C. Paulino, *Elife* **2019**, 8, e44365.
[77] J. Biber, N. Hernando, I. Forster, *Annu. Rev. Physiol.* **2013**, 75, 535-550.
[78] D. B. Rhoads, F. B. Waters, W. Epstein, *J. Gen. Physiol.* **1976**, 67, 325-341.
[79] M. Diskowski, A. R. Mehdipour, D. Wunnicke, D. J. Mills, V. Mikusevic, N. Bärland, J. Hoffmann, N. Morgner, H.-J. Steinhoff, G. Hummer, J. Vonck, I. Hänelt, *Elife* **2017**, 6, e24303.
[80] S. Okumoto, A. Jones, W. B. Frommer, *Annu. Rev. Plant Biol.* **2012**, 63, 663-706.
[81] L. Lindenburg, M. Merkx, *Sensors* **2014**, 14, 11691-11713.

[82] H. Imamura, K. P. H. Nhat, H. Togawa, K. Saito, R. Iino, Y. Kato-Yamada, T. Nagai, H. Noji, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, 106, 15651-15656.
[83] H. Yaginuma, S. Kawai, K. V Tabata, K. Tomiyama, A. Kakizuka, T. Komatsuzaki, H. Noji, H. Imamura, *Sci. Rep.* **2014**, 4, 6522.
[84] M. Tantama, J. R. Martínez-François, R. Mongeon, G. Yellen, *Nat. Commun.* **2013**, 4, 2550.
[85] M. A. Lobas, R. Tao, J. Nagai, M. T. Kronschläger, P. M. Borden, J. S. Marvin, L. L. Looger, B. S. Khakh, *Nat. Commun.* **2019**, 10, 711.
[86] Y. Zhao, Q. Hu, F. Cheng, N. Su, A. Wang, Y. Zou, H. Hu, X. Chen, H.-M. Zhou, X. Huang, K. Yang, Q. Zhu, X. Wang, J. Yi, L. Zhu, X. Qian, L. Chen, Y. Tang, J. Loscalzo, Y. Yang, *Cell Metab.* **2015**, 21, 777-789.
[87] R. Tao, Y. Zhao, H. Chu, A. Wang, J. Zhu, X. Chen, Y. Zou, M. Shi, R. Liu, N. Su, J. Du, H. M. Zhou, L. Zhu, X. Qian, H. Liu, J. Loscalzo, Y. Yang, *Nat. Methods* **2017**, 14, 720-728.
[88] G. Miesenböck, D. A. De Angelis, J. E. Rothman, *Nature* **1998**, 394, 192-195.
[89] M. Tantama, Y. P. Hung, G. Yellen, *J. Am. Chem. Soc.* **2011**, 133, 10034-10037.
[90] K. Kano, J. H. Fendler, *Biochim. Biophys. Acta* **1978**, 509, 289-299.
[91] M. R. James-Kracke, *J. Cell. Physiol.* **1992**, 151, 596-603.
[92] A. P. Singh, P. Nicholls, *J. Biochem. Biophys. Methods* **1985**, 11, 95-108.
[93] B. Liu, B. Poolman, A. J. Boersma, *ACS Chem. Biol.* **2017**, 12, 2510-2514.
[94] A. J. Boersma, I. S. Zuhorn, B. Poolman, *Nat. Methods* **2015**, 12, 227-229.
[95] B. Liu, C. Åberg, F. J. van Eerden, S. J. Marrink, B. Poolman, A. J. Boersma, *Biophys. J.* **2017**, 112, 1929-1939.
[96] D. Gnutt, M. Gao, O. Brylski, M. Heyden, S. Ebbinghaus, *Angew. Chem. Int. Ed. Engl.* **2015**, 54, 2548-2551.
[97] M. K. Kuimova, *Phys. Chem. Chem. Phys.* **2012**, 14, 12671-12686.
[98] S.-C. Lee, J. Heo, H. C. Woo, J.-A. Lee, Y. H. Seo, C.-L. Lee, S. Kim, O.-P. Kwon, *Chemistry* **2018**, 24, 13706-13718.
[99] Y. Shen, S.-Y. Wu, V. Rancic, A. Aggarwal, Y. Qian, S.-I. Miyashita, K. Ballanyi, R. E. Campbell, M. Dong, *Commun. Biol.* **2019**, 2, 1-10.
[100] A. M. Feist, C. S. Henry, J. L. Reed, M. Krummenacker, A. R. Joyce, P. D. Karp, L. J. Broadbelt, V. Hatzimanikatis, B. Ø. Palsson, *Mol. Syst. Biol.* **2007**, 3, 121.
[101] M. S. Rappé, S. A. Connon, K. L. Vergin, S. J. Giovannoni, *Nature* **2002**, 418, 630-633.
[102] P. Mateos-Gil, P. Tarazona, M. Vélez, *FEMS Microbiol. Rev.* **2019**, 43, 73-87.
[103] B. H. Kim, G. M. Gadd, *Bacterial Physiology and Metabolism*, Cambridge University Press, **2008**.

Entry for the Table of Contents (Please choose one layout)

Layout 1:

MINIREVIEW

Text for Table of Contents

((Insert TOC Graphic here))

*Author(s), Corresponding Author(s)**

Page No. – Page No.

Title

Layout 2:

MINIREVIEW

((Insert TOC Graphic here))

*Author(s), Corresponding Author(s)**

Page No. – Page No.

Title

Text for Table of Contents